On page 1, beginning at line 1, the replacement paragraph to read:

SYSTEM FOR DISTRIBUTING LOAD OVER MULTIPLE SERVERS AT AN

INTERNET SITE

On page 1, beginning at line 5, the replacement paragraph to read:

This application is a continuation in part of co-pending application Serial No.

08/552,807, which is incorporated herein by reference for all purposes.

On page 1, beginning at line 8, the replacement paragraph to read:

This application is related to co-pending application Serial Nos. 08/850,730 and

08/850,836, filed concurrently herewith and subsequently issued as U.S. Pat. No. 6,061,349

issued May 9, 2000 and No. 6,104,717 issued August 15, 2000, respectively, which are

incorporated herein by reference for all purposes.

On page 12, beginning on line 19, the replacement paragraph to read as follows:

Requests to internet site 100 from external sites on Internet 102 are routed through Local

Director 110. Local Director 110 determines which server of group of TCP based servers 112

should receive the request. Briefly, it does this as follows. A virtual IP address is defined for

internet site 100. This virtual IP address is the IP address which the outside world, including the

rest of the Internet 102, uses to access internet site 100. To an outside client, it appears that a

single virtual machine having a single virtual IP address services internet site 100. The

individual identities and IP addresses of the individual servers within the group of TCP based

servers 112 are not evident to the user. In certain embodiments, group of TCP based servers 112

may implement more than ● virtual server as described in U.S. Pa● No. 6,061,349, which is incorporated herein by reference. In such embodiments, a plurality of virtual machines are implemented on different port numbers of a set of real or physical machines. In accordance with the present invention, each virtual machine may allocate connections to a plurality of physical machines.

On page 14, beginning at line 8, the replacement paragraph to read as follows:

Local Director 110 thus operates to distribute packets among group of TCP based servers 112 by intercepting each packet sent to a virtual machine at internet site 100 and changing the destination IP address in the packet from a virtual IP address which corresponds to all of internet site 100 to a real IP address which corresponds to a single physical machine located at internet site 100. In certain embodiments, Local Director 110 includes more than one virtual machine IP address and therefore routes connections for more than one virtual machine to a set of physical machines through the physical machine's ports. Additionally, in such embodiments, it is also possible that each physical machine is mapped to more than one virtual machine. Such a system is described in detail in co-pending application Serial No. 08/850,730, now U.S. Patent No. 6,061,349, previously incorporated by reference.

On page 15, beginning at line 1, the replacement paragraph to read:

By adopting a single virtual IP address for the entire server group of TCP based servers 112, the problems of round robin DNS and DNS caching are avoided. Specifically, any connection made to the virtual IP address of a virtual machine is perceived by the connecting entity as a connection to the virtual machine and not as a connection to the physical machine to which the connection is physically made. The connecting entity never discovers the real IP address of the real machine handling the connection since, for outgoing packets, the real machine

source IP address is replaced with the virtual machine IP address by Local Director. Therefore a connecting entity which caches IP addresses using DNS caching caches the virtual machine IP address and will not address connection requests exclusively to any one server from group of TCP based servers 112 to the exclusion of the other servers in group of TCP based servers 112.

On page 17, beginning at line 15, the replacement paragraph to read:

In certain embodiments, servers are failed when they fail to make a certain number of consecutive connections corresponding to a failure threshold. In some embodiments, each failed connection itself fails only after repeated attempts to make the connection are unsuccessful. The predicted response time for such a machine would still match the aged predicted response time from its last successful response. That predicted response time would be unduly optimistic since if the server has failed, then, in fact, the actual response time is going to be at least as long as it takes to fix the machine and bring it back up on line. The selection of the server based on its unrealistically good response time is therefore overridden by a failure flag. A further description of a system in which failures of individual physical machines are determined and failed machines are tested to determine if they can be placed on line again is described in detail in co-pending application Ser. No. 08/850,836, now U.S. Pat. No. 6,104,717, filed concurrently herewith, which is incorporated herein by reference for all purposes.